Stacked RIS-Assisted Dual-Polarized UAV-RSMA Networks

Yifu Sun[†], Yonggang Zhu[†], Haotong Cao[‡], Zhi Lin[§], Kang An[†], Feng Tian[◁], Kai-Kit Wong[▷], and Jiangzhou Wang[◇]
[†] The Sixty-third Research Institute, National University of Defense Technology, Nanjing, China
[‡] Department of Computing, Hong Kong Polytechnic University, Hong Kong SAR, China
[§] College of Electronic Engineering, National University of Defense Technology, Hefei, China
[◁] Collegue of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China
[▷] Department of Electronic and Electrical Engineering, University College London, London, U.K.
[◇] School of Engineering, University of Kent, Canterbury, U.K.
Email: sunyifu_nudt@163.com, zhumaka1982@163.com, haotong.cao@polyu.edu.hk, linzhi945@163.com, ankang89@nudt.edu.cn, tianf@njupt.edu.cn, kai-kit.wong@ucl.ac.uk, j.z.wang@kent.ac.uk

***Abstract*—Due to the users' overlapping channels and the open nature of the wireless medium, inter-user interference and malicious jamming attacks deteriorate the performance of unmanned aerial vehicle (UAV) communications. With this focus, this paper proposes a novel integration of dual polarization, rate-splitting multiple access (RSMA), and stacked reconfigurable intelligent surface (RIS) transceiver into UAV networks, thus simultaneously mitigating the inter-user interference and malicious interference by fully exploiting their potentials in the power, space, and polarization domains. Building upon this architectural framework, a generalized sum rate maximization problem is formulated under the jammer's imperfect angular channel state information and unknown cross-polarization discrimination. To efficiently tackle the challenges posed by the intractable non-convex design problem with both high-dimensional variables and the multiple QoS constraints, a low-complexity optimization framework is presented, where a discretization method combined with quadratic property, a reduced-majorization-minimization algorithm, and two computationally efficient algorithms using block successive upper-bound minimization are developed to obtain the semi-closed-form solutions. Finally, numerical simulations verify the superiority and validity of our proposed architecture and optimization framework over benchmarks.**

***Index Terms*—UAV communications, RSMA, stacked RIS, dual polarization, interference mitigation, jamming suppression.**

I. Introduction

UNMANNED aerial vehicle (UAV) has emerged as indispensable component of sixth-generation (6G) networks due to its high maneuverability and low cost [1]. However, owing to the inherent broadcast and superposition properties of wireless media, both the inter-user interference and malicious jamming attacks pose severe threats to UAV communications [2]. To cope with the abovementioned issues, various conventional techniques have been proposed in the literature, such as multiple access techniques for interference mitigation [3], frequency hopping technique for jamming suppression [4], and massive multiple-input multiple-output (MIMO) technique for both [5]. Nevertheless, conventional multiple access techniques lead to either insufficient radio resources utilization or extra spectral resources, and MIMO technique is limited by channel state information (CSI) imperfections and the degrees of freedom (DoFs) due to the limited array aperture size [6].

Fortunately, rate-splitting multiple access (RSMA) and reconfigurable intelligent surface (RIS) have recently emerged as promising paradigms for overcoming the limitations of convetional multiple access schemes and facilitating the employment of large-scale arrays, respectively, which promises to alleviate the inter-user interference and malicious jamming issues [3], [5]. The benefits of RSMA have been investigated thoroughly, such as delivering high spectral/energy efficiency [7], improving DoFs [8], and enhancing the robustness to CSI imperfection/high mobility/eavesdropping attacks [9]. However, the abovementioned RSMA schemes rely on the delicate SIC technique, which limits its applications. This fact calls for the effective strategies for handling the adverse effects of SIC imperfections in RSMA.

Current state-of-art for RIS-aided wireless communications can be roughly categorized into two main streams, i.e., RIS-based passive reflector [10] and RIS-aided active transceiver [5], [11]. The former RIS architecture has been widely proposed in UAV networks for modifying the propagation environment [12], while the latter one has been proposed to address the dilemma of the prohibitive hardware cost in deploying the large-scale arrays at the transceiver [5], [11]. However, the utilization of RIS-reflector in UAV networks suffers from the extremely severe "double fading" effects [11] and high mobility [12]. Fortunately, RIS-aided active transceiver has been proven to effectively overcome the "double fading" effects and mobility [13], which includes single-layer [13] and stacked architectures [5], [14]. Compared to single-layer one, stacked RIS-transceiver offers a more desirable miniaturization, and the amplitude of refracted signal though RIS can be partially controlled in a larger range [14].

Recently, polarization has been exploited to complement the RSMA's SIC imperfection, and further increase the number of RIS units without increasing the array aperture [15]. Particularly, the dual-polarized RSMA provided a platform where all the receivers can decode the common and private signals simultaneously in two completely orthogonal polarizations without SIC, thus removing the inherent SIC imperfection in RSMA. In addition, by stacking two single-polarized RIS in two orthogonal polarizations together, the number of RIS units in the dual-polarized RIS doubles that of the single-polarized one within the same array aperture, thereby doubling the channel capacity.

Motivated by these studies above, this paper proposes a novel integration of dual-polarized RSMA and stacked RIS-aided transceiver into the UAV networks, where dual polarization can naturally remove the inherent SIC imperfection in RSMA and enlarge the DoFs of stacked RIS without increasing the array aperture. Besides, a worst-case sum achievable rate maximization problem is formulated under the jammer's imperfect angular CSI and unknown cross-polarization discrimination (XPD). To handle the formulated intractable problem, a low-complexity optimization framework is proposed by leveraging

the discretization method, properties of the quadratic function, educed-majorization-minimization (R-MM) algorithm, and block successive upper-bound minimization (BSUM), which admits the semi-closed-form solutions. Finally, numerical simulations demonstrate that the proposed architecture and optimization framework are superior to the state-of-the-art schemes.

II. System Model and Problem Formulation

A. Stacked RIS-Aided Dual-Polarized UAV-RSMA Networks

Fig. 1 illustrates the considered RIS-assisted UAV-RSMA networks, where a UAV is equipped with a dual-polarized stacked RIS-assisted transmitter, and thus assists to establish the reliable communication links with K dual-polarized stacked RIS-aided receivers/users in the presence of a malicious jammer. More specifically, the UAV's RIS-transmitter is equipped with $N_{\mathrm{L}}{=}N_{\mathrm{L}1} \times N_{\mathrm{L}2}$ pairs of dual-polarized Tx antennas and A layers of RIS having $M_{\mathrm{L},a}{=}M_{\mathrm{L},a1} \times M_{\mathrm{L},a2}$ dual-polarized units on a-th layer, while k-th user's RIS-receiver is equipped with $N_{\mathrm{U},k}{=}N_{\mathrm{U},k1} \times N_{\mathrm{U},k2}$ pairs of dual-polarized Rx antennas and R layers of RIS having $M_{\mathrm{U},kr} = M_{\mathrm{U},kr1} \times M_{\mathrm{U},kr2}$ dual-polarized units on r-th layer. In addition, we assume that the jammer adopts one pair of dual-polarized omnidirectional Tx antennas to impair the users' signal reception at both the vertical and the horizontal polarization. To effectively mitigate the inter-user interference, RSMA is embedded into the dual-polarized transceiver in this paper. To simplify the presentation, we denote the stacked RIS in the RIS-aided transmitter and receiver as the stacked T-RIS and R-RIS, respectively. Thus, the dual-polarized channels between UAV's Tx antennas and the first layer T-RIS, between the a-th layer T-RIS and the $(a+1)$-th layer T-RIS, between UAV and k-th user, between kr-th R-RIS and $k(r+1)$-th R-RIS, between kR-th R-RIS and k-th user's Rx antennas, and between the jammer and k-th user are denoted as $\mathbf{B}_{\mathrm{L},1} \in \mathbb{C}^{2M_{\mathrm{L},1}\times 2N_{\mathrm{L}}}$, $\mathbf{B}_{\mathrm{L},a} \in \mathbb{C}^{2M_{\mathrm{L},a}\times 2M_{\mathrm{L},a}}$, $\mathbf{D}_{\mathrm{U},k} \in \mathbb{C}^{2M_{\mathrm{U},k1}\times 2M_{\mathrm{L},A}}$, $\mathbf{B}_{\mathrm{U},kr} \in \mathbb{C}^{2M_{\mathrm{U},kr}\times 2M_{\mathrm{U},kr}}$, $\mathbf{B}_{\mathrm{U},kR} \in \mathbb{C}^{2N_{\mathrm{U},k}\times 2M_{\mathrm{U},kR}}$, and $\mathbf{D}_{\mathrm{J},k} \in \mathbb{C}^{2M_{\mathrm{U},k1}\times 2}$, respectively. Here, kr-th R-RIS denotes the r-th layer R-RIS in the k-th user. Under the dual-polarized settings, the involved channels can be partitioned as

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}^{\mathrm{vv}} & \mathbf{D}^{\mathrm{vh}} \\ \mathbf{D}^{\mathrm{hv}} & \mathbf{D}^{\mathrm{hh}} \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} \mathbf{B}^{\mathrm{vv}} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{\mathrm{hh}} \end{bmatrix}, \tag{1}$$

where $\mathbf{H}^{ij}$ is a channel coefficient from the polarization i to the polarization j, where $i,j \in \{\mathrm{v},\mathrm{h}\}$, with v and h denoting the vertical and horizontal polarization, respectively. Moreover, they can be further expressed as [15]

$$\mathbf{H}{=}\begin{bmatrix} \mathbf{H}^{\mathrm{vv}} & \mathbf{H}^{\mathrm{vh}} \\ \mathbf{H}^{\mathrm{hv}} & \mathbf{H}^{\mathrm{hh}} \end{bmatrix}{=}\begin{bmatrix} \sqrt{1-\alpha}e^{j\varphi^{\mathrm{vv}}}\widehat{\mathbf{H}} & \sqrt{\alpha}e^{j\varphi^{\mathrm{vh}}}\widehat{\mathbf{H}} \\ \sqrt{\alpha}e^{j\varphi^{\mathrm{hv}}}\widehat{\mathbf{H}} & \sqrt{1-\alpha}e^{j\varphi^{\mathrm{hh}}}\widehat{\mathbf{H}} \end{bmatrix}, \tag{2}$$

where $0 \le \alpha \le 1$ is the inverse XPD which measures the proportion of radiated power converted from polarization v to polarization h and vice versa. Besides, φ^{ij} are the random phase shifts in the corresponding polarization ij. Note that due to the extreme short distance between the adjacent layer in the stacked RIS-transmitter/receiver, XPD does not exist in $\mathbf{B}$.

B. Signal Transmission Model

Denote s_{C} and $s_{\mathrm{P},k}$ as the common symbol for all users and the private symbols for specific user k, respectively. Prior to transmission, s_{C} and $s_{\mathrm{P},k}$ are processed by the digital

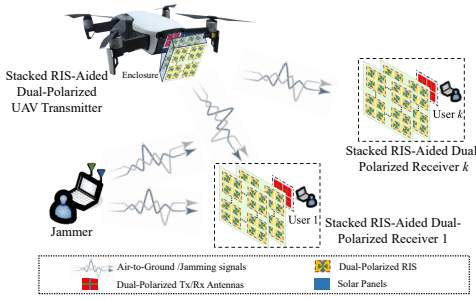


Fig. 1: System model.

precoder $\mathbf{w}_{\mathrm{C}} = [\mathbf{w}_{\mathrm{C}}^{\mathrm{v}}; \mathbf{0}_{N_{\mathrm{L}}\times 1}]$ and $\mathbf{w}_{\mathrm{P},k} = \left[\mathbf{0}_{N_{\mathrm{L}}\times 1}; \mathbf{w}_{\mathrm{P},k}^{\mathrm{h}}\right]$, respectively, where $\mathbf{w}_{\mathrm{C}}^{\mathrm{v}}$ and $\mathbf{w}_{\mathrm{P},k}^{\mathrm{h}}$ are the precoder for the common and private symbols in the polarization v and h. Hence, the digital beamformer transmitted to stacked T-RIS is $\mathbf{w}_{\mathrm{C}}s_{\mathrm{C}} + \sum_{k=1}^{K}\mathbf{w}_{\mathrm{P},k}s_{\mathrm{P},k}$. Then, the dual-polarized phase-shifter pairs on T-RIS can impose phase shifts to the incident signals of vertical and horizontal polarizations, and then forward them to the dual-polarized transmit units via the microstrip. Here, we define the phase-shift matrix at the a-th layer T-RIS as $\mathbf{P}_{\mathrm{L},a}{=}\,\mathrm{Blkdiag}\left\{\mathbf{P}_{\mathrm{L},a}^{\mathrm{vv}}, \mathbf{P}_{\mathrm{L},a}^{\mathrm{hh}}\right\} \in \mathbb{C}^{2M_{\mathrm{L},a}\times 2M_{\mathrm{L},a}}$, where $\mathbf{P}_{\mathrm{L},a}^{ii} = \mathrm{diag}\left\{e^{j\phi_{\mathrm{L},a1}^{ii}}, \cdots, e^{j\phi_{\mathrm{L},aM_{\mathrm{L},a}}^{ii}}\right\} \in \mathbb{C}^{M_{\mathrm{L},a}\times M_{\mathrm{L},a}}$ is the phase-shift matrix on polarization $i \in \{\mathrm{v},\mathrm{h}\}$. Thus, the hybrid digital-analog beamformer transmitted to the users is

$$\mathbf{x} = \prod_{a=1}^{A}\mathbf{P}_{\mathrm{L},a}\mathbf{B}_{\mathrm{L},a}\left(\mathbf{w}_{\mathrm{C}}s_{\mathrm{C}} + \sum_{m=1}^{K}\mathbf{w}_{\mathrm{P},m}s_{\mathrm{P},m}\right). \tag{3}$$

Meanwhile, the jammer divides its total jamming power into two parts in vertical and horizontal polarization for interrupting the legitimate transmission, i.e., $\mathbf{x}_{\mathrm{J}} = [\sqrt{p_{\mathrm{J}}^{\mathrm{v}}}s_{\mathrm{J}}^{\mathrm{v}}; \sqrt{p_{\mathrm{J}}^{\mathrm{h}}}s_{\mathrm{J}}^{\mathrm{h}}]$.

Then, the signals $\mathbf{x}$ and $\mathbf{x}_{\mathrm{J}}$ are superimposed at the k-th stacked RIS-receiver via the far-field channels $\mathbf{D}_{\mathrm{U},k}$ and $\mathbf{D}_{\mathrm{J},k}$. Here, the phase-shift matrix of the kr-th R-RIS is defined as $\mathbf{P}_{\mathrm{U},kr} = \mathrm{Blkdiag}\left\{\mathbf{P}_{\mathrm{U},kr}^{\mathrm{vv}}, \mathbf{P}_{\mathrm{U},kr}^{\mathrm{hh}}\right\} \in \mathbb{C}^{2M_{\mathrm{U},kr}\times 2M_{\mathrm{U},kr}}$, and the digital decoder is represented as $\mathbf{V}_k = \mathrm{Blkdiag}\left\{\mathbf{v}_k^{\mathrm{v}}, \mathbf{v}_k^{\mathrm{h}}\right\} \in {}^{2N_{\mathrm{U},k}\times 2}$, where $\mathbf{P}_{\mathrm{U},kr}^{ii} = \mathrm{diag}\left\{e^{j\phi_{\mathrm{U},kr1}^{ii}}, \cdots, e^{j\phi_{\mathrm{U},krM_{\mathrm{U},kr}}^{ii}}\right\} \in \mathbb{C}^{M_{\mathrm{U},kr}\times M_{\mathrm{U},kr}}$ and $\mathbf{v}_k^i$ is the phase-shift matrix and digital decoder on polarization $i \in \{\mathrm{v},\mathrm{h}\}$, respectively. Hence, the received signal at the k-th user is given by

$$\mathbf{y}_k = \widetilde{\mathbf{V}}^H\left(\mathbf{D}_{\mathrm{U},k}\mathbf{x} + \mathbf{D}_{\mathrm{J},k}\mathbf{x}_{\mathrm{J}}\right) + \widetilde{\mathbf{n}}_k, \tag{4}$$

where $\widetilde{\mathbf{n}}_k = [\mathbf{v}_k^{\mathrm{v},H}\mathbf{n}_k^{\mathrm{v}}; \mathbf{v}_k^{\mathrm{h},H}\mathbf{n}_k^{\mathrm{h}}]$, $\mathbf{n}_k^i \sim \mathcal{CN}\left(\mathbf{0}_{N_{\mathrm{U},k}}, \sigma^2\mathbf{I}_{N_{\mathrm{U},k}}\right)$ is the additive white noise vector in polarization $i \in \{\mathrm{v},\mathrm{h}\}$, and $\widetilde{\mathbf{V}}^H = \mathbf{V}^H\prod_{r=1}^{R}\mathbf{B}_{\mathrm{U},kr}\mathbf{P}_{\mathrm{U},kr}$. Thus, the information rates to decode s_{C} and $s_{\mathrm{P},k}$ at the k-th user are expressed as

$$\begin{aligned} R_{\mathrm{C},k}{=}\log\left(1{+}\mathbf{v}_k^{\mathrm{v},H}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv}}\mathbf{w}_{\mathrm{C}}^{\mathrm{v}}\mathbf{w}_{\mathrm{C}}^{\mathrm{v},H}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv},H}\mathbf{v}_k^{\mathrm{v}}\left(\mathbf{v}_k^{\mathrm{v},H}\mathbf{C}_{\mathrm{U},k}^{\mathrm{v}}\mathbf{v}_k^{\mathrm{v}}\right)^{-1}\right), \\ R_{\mathrm{P},k}{=}\log\left(1{+}\mathbf{v}_k^{\mathrm{h},H}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh}}\mathbf{w}_{\mathrm{P},k}^{\mathrm{h}}\mathbf{w}_{\mathrm{P},k}^{\mathrm{h},H}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh},H}\mathbf{v}_k^{\mathrm{h}}\left(\mathbf{v}_k^{\mathrm{h},H}\mathbf{C}_{\mathrm{U},k}^{\mathrm{h}}\mathbf{v}_k^{\mathrm{h}}\right)^{-1}\right), \end{aligned} \tag{5}$$

where

$$\mathbf{C}_{\mathrm{U},k}^{\mathrm{v}}=\mathbf{C}_{\mathrm{U}1,k}^{\mathrm{v}}+\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{vv}}\left(p_{\mathrm{J}}^{\mathrm{v}}\mathbf{d}_{\mathrm{J},k}^{\mathrm{vv}}\mathbf{d}_{\mathrm{J},k}^{\mathrm{vv},H}+p_{\mathrm{J}}^{\mathrm{h}}\mathbf{d}_{\mathrm{J},k}^{\mathrm{vh}}\mathbf{d}_{\mathrm{J},k}^{\mathrm{vh},H}\right)\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{vv},H},$$

$$\mathbf{C}_{\mathrm{U},k}^{\mathrm{h}}=\mathbf{C}_{\mathrm{U}1,k}^{\mathrm{h}}+\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{hh}}\left(p_{\mathrm{J}}^{\mathrm{h}}\mathbf{d}_{\mathrm{J},k}^{\mathrm{hh}}\mathbf{d}_{\mathrm{J},k}^{\mathrm{hh},H}+p_{\mathrm{J}}^{\mathrm{v}}\mathbf{d}_{\mathrm{J},k}^{\mathrm{hv}}\mathbf{d}_{\mathrm{J},k}^{\mathrm{hv},H}\right)\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{hh},H},$$

$$\mathbf{C}_{\mathrm{U}1,k}^{\mathrm{v}}=\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vh}}\left(\sum\nolimits_{m=1}^{K}\mathbf{w}_{\mathrm{P},m}^{\mathrm{h}}\mathbf{w}_{\mathrm{P},m}^{\mathrm{h},H}\right)\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vh},H}+\sigma_{k}^{\mathrm{v},2}\mathbf{I}_{N_{\mathrm{U},k}},$$

$$\mathbf{C}_{\mathrm{U}1,k}^{\mathrm{h}}=\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh}}\left(\sum\nolimits_{m\neq k}^{K}\mathbf{w}_{\mathrm{P},m}^{\mathrm{h}}\mathbf{w}_{\mathrm{P},m}^{\mathrm{h},H}\right)\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh},H}+\sigma_{k}^{\mathrm{h},2}\mathbf{I}_{N_{\mathrm{U},k}}+\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hv}}\mathbf{w}_{\mathrm{C}}^{\mathrm{v}}\mathbf{w}_{\mathrm{C}}^{\mathrm{v},H}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hv},H},$$

$$\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vh}}=\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{vv}}\mathbf{D}_{\mathrm{U},k}^{\mathrm{vh}}\mathbf{\Omega}_{\mathrm{L}}^{\mathrm{hh}},\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hv}}=\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{hh}}\mathbf{D}_{\mathrm{U},k}^{\mathrm{hv}}\mathbf{\Omega}_{\mathrm{L}}^{\mathrm{vv}}$$

$$\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh}}=\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{hh}}\mathbf{D}_{\mathrm{U},k}^{\mathrm{hh}}\mathbf{\Omega}_{\mathrm{L}}^{\mathrm{hh}},\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv}}=\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{vv}}\mathbf{D}_{\mathrm{U},k}^{\mathrm{vv}}\mathbf{\Omega}_{\mathrm{L}}^{\mathrm{vv}}.$$

$$\mathbf{\Omega}_{\mathrm{L}}^{\mathrm{vv}}=\prod\nolimits_{a=1}^{A}\mathbf{P}_{\mathrm{L},a}^{\mathrm{vv}}\mathbf{B}_{\mathrm{L},a}^{\mathrm{vv}},\mathbf{\Omega}_{\mathrm{L}}^{\mathrm{hh}}=\prod\nolimits_{a=1}^{A}\mathbf{P}_{\mathrm{L},a}^{\mathrm{hh}}\mathbf{B}_{\mathrm{L},a}^{\mathrm{hh}},$$

$$\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{vv}}=\prod\nolimits_{r=1}^{R}\mathbf{B}_{\mathrm{U},kr}^{\mathrm{vv}}\mathbf{P}_{\mathrm{U},kr}^{\mathrm{vv}},\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{hh}}=\prod\nolimits_{r=1}^{R}\mathbf{B}_{\mathrm{U},kr}^{\mathrm{hh}}\mathbf{P}_{\mathrm{U},kr}^{\mathrm{hh}}.$$

Since s_{C} is decoded by all the users, the information rate for common message is defined as $R_{\mathrm{C}}=\min\{R_{\mathrm{C},1},\cdots,R_{\mathrm{C},K}\}$, which can guarantee that all users can successfully decode the common symbol s_{C} [3]. As R_{C} is shared by all users, where the k-th user is allocated with a portion C_k corresponding to common part $M_{\mathrm{C},k}$, and we have $R_{\mathrm{C}}=\sum_{k=1}^{K}C_k$. Thus, the total information rate for the k-th user is $C_k+R_{\mathrm{P},k}$.

C. Problem Formulation

Due to the fact that the jammers are not expected to cooperate with the user for the channel estimation [16], the illegitimate XPD $\alpha_{\mathrm{J},k}$ and CSI $\widehat{\mathbf{d}}_{\mathrm{J},k}$ are challenging to be obtained. To account for the harmful effects on the system performance, we assume that the jammer's CSI $\widehat{\mathbf{d}}_{\mathrm{J},k}$ belongs to a given angular range Δ and its XPD $\alpha_{\mathrm{J},k}$ belongs to an unknown set $\bigcirc$, which is

$$\begin{aligned}\Delta&=\left\{\widehat{\mathbf{d}}_{\mathrm{J},k}\middle|\,\theta\in[\theta_L,\theta_U],\varphi\in[\varphi_L,\varphi_U],\forall k\right\},\\ \bigcirc&=\left\{\underset{\sim}{\bigcirc}\alpha_{\mathrm{J},k}\middle|\,\alpha_{\mathrm{J},k}\in[0,1],\forall k\right\},\end{aligned}\tag{6}$$

where $\underset{\sim}{\bigcirc}\alpha$ denotes α is an unknown value, θ_{U} and θ_{L} denote the upper and lower bounds of azimuth angle, φ_{U} and φ_{L} are the upper and lower bounds of elevation angle. Next, we formulate a general worst-case sum achievable rate maximization problem:

$$\begin{aligned}\max_{\mathbf{V}_k,\mathbf{w}_{\mathrm{C}}^{\mathrm{v}},\mathbf{w}_{\mathrm{P},k}^{\mathrm{h}},\mathbf{P}_{\mathrm{L},a},\mathbf{c},\mathbf{P}_{\mathrm{U},kr}}\ &\min_{\bigcirc,\Delta}\sum\nolimits_{k=1}^{K}C_k+R_{\mathrm{P},k}\\ s.t.\ \mathrm{C1}:&\min_{\bigcirc,\Delta}C_k+R_{\mathrm{P},k}\geq R_{k,\min},\forall k,\\ \mathrm{C2}:&\min_{\bigcirc,\Delta}R_{\mathrm{C},k}\geq\sum\nolimits_{q=1}^{K}C_q,\forall k,\\ \mathrm{C3}:&\left\|\mathbf{w}_{\mathrm{C}}^{\mathrm{v}}\right\|^2+\sum\nolimits_{k=1}^{K}\left\|\mathbf{w}_{\mathrm{P},k}^{\mathrm{h}}\right\|^2\leq P_{\max},\\ \mathrm{C4}:&\left|[\mathbf{P}_{\mathrm{L},a}]_{n,n}\right|=1,\left|[\mathbf{P}_{\mathrm{U},kr}]_{n,n}\right|=1,\forall n,k,r,a,\end{aligned}\tag{7}$$

where $\mathbf{c}=(C_1,\cdots,C_K)^T$.

III. Efficient Optimization Framework for (7)

A. Heuristic Robust Decoder Design for $\mathbf{V}_k$

First, we focus on investigating the design of $\mathbf{V}_k$. According to [11], the linear minimum-mean-square-error (MMSE) detector is the optimal solution for $\mathbf{V}_k$, which maximizes the signal-to-interference plus-noise-ratio (SINR), which is expressed as

$$\mathbf{v}_k^{\mathrm{v}}=\mathbf{C}_{\mathrm{U},k}^{\mathrm{v},-1}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv}}\mathbf{w}_{\mathrm{C}}^{\mathrm{v}},\ \mathbf{v}_k^{\mathrm{h}}=\mathbf{C}_{\mathrm{U},k}^{\mathrm{h},-1}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh}}\mathbf{w}_{\mathrm{P},k}^{\mathrm{h}},\tag{8}$$

where $\mathbf{C}_{\mathrm{U},k}^{\mathrm{v}}$, $\mathbf{C}_{\mathrm{U},k}^{\mathrm{h}}$, $\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv}}$, and $\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh}}$ are given in (5). However, due to the imperfect angular CSI Δ and unknown XPD set $\bigcirc$, i.e., the term of $\mathbf{D}_{\mathrm{J},k}$ inside $\mathbf{C}_{\mathrm{U},k}^{\mathrm{v}}$ and $\mathbf{C}_{\mathrm{U},k}^{\mathrm{h}}$ cannot be obtained, the solution of $\mathbf{V}_k$ in (12) is infeasible. To make (12) feasible, the following propositions are proposed to handle the CSI imperfection Δ and unknown XPD set $\bigcirc$.

Proposition 1 (Discretization Method for Δ): After uniformly discretizing all the angles inside Δ, i.e.,

$$\begin{aligned}\theta^{(p)}&=\theta_L+(i-1)\,\Delta\theta,\ p=1,\cdots,Q_1,\\ \varphi^{(q)}&=\varphi_L+(j-1)\,\Delta\varphi,\ q=1,\cdots,Q_2,\end{aligned}\tag{9}$$

where Q_1 and Q_2 are the number of samples of θ and φ, $\Delta\theta=(\theta_U-\theta_L)/(Q_1-1)$, and $\Delta\varphi=(\varphi_U-\varphi_L)/(Q_2-1)$, the worst-case CSI of $\widehat{\mathbf{d}}_{\mathrm{J},k}$ inside $\mathbf{D}_{\mathrm{J},k}$, i.e., $\widetilde{\mathbf{d}}_{\mathrm{J},k}\widetilde{\mathbf{d}}_{\mathrm{J},k}^{H}$, is

$$\widetilde{\mathbf{d}}_{\mathrm{J},k}\widetilde{\mathbf{d}}_{\mathrm{J},k}^{H}=\sum\nolimits_{p=1}^{Q_1}\sum\nolimits_{q=1}^{Q_2}\left(1/Q_1Q_2\right)\widehat{\mathbf{d}}_{\mathrm{J},k}^{(p,q)}\widehat{\mathbf{d}}_{\mathrm{J},k}^{(p,q),H},\tag{10}$$

where $\widehat{\mathbf{d}}_{\mathrm{J},k}^{(p,q)}$ is the selected element of $\left\{\theta^{(p)},\varphi^{(q)}\right\}$.

Proof: Please refer to [11]. ■

Proposition 2: After safely ignoring the jamming channels' random phase shifts, i.e., $\phi_{\mathrm{J},k}^{ij},ij\in\{\mathrm{vv},\mathrm{vh},\mathrm{hv},\mathrm{hh}\}$ inside $\mathbf{D}_{\mathrm{J},k}$, we can obtain the worst-case XPD $\widetilde{\alpha}_{\mathrm{J},k}$ insides $\bigcirc$, i.e.,

$$\begin{cases}\widetilde{\alpha}_{\mathrm{J},k}=1/2,p_{\mathrm{J}}^{\mathrm{v}}=p_{\mathrm{J}}^{\mathrm{h}},\\ \widetilde{\alpha}_{\mathrm{J},k}=\underset{\alpha_{\mathrm{J},k}\in[0,1]}{\arg\min}\frac{-\overline{p}^2\alpha_{\mathrm{J},k}^2+\overline{z}_k\alpha_{\mathrm{J},k}+\overline{d}_k}{-\overline{p}^2\alpha_{\mathrm{J},k}^2+\widetilde{z}_k\alpha_{\mathrm{J},k}+\widetilde{d}_k},p_{\mathrm{J}}^{\mathrm{v}}\neq p_{\mathrm{J}}^{\mathrm{h}},\end{cases}\tag{11}$$

where $\overline{d}_k=\overline{z}_{\mathrm{C},k}\overline{z}_{\mathrm{P},k},\widetilde{d}_k=\widetilde{z}_{\mathrm{C},k}\widetilde{z}_{\mathrm{P},k}$,

$$\begin{aligned}a_{\mathrm{C},k}&=\mathbf{v}_k^{\mathrm{v},H}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv}}\mathbf{w}_{\mathrm{C}}^{\mathrm{v}}\mathbf{w}_{\mathrm{C}}^{\mathrm{v},H}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv},H}\mathbf{v}_k^{\mathrm{v}},\\ a_{\mathrm{P},k}&=\mathbf{v}_k^{\mathrm{h},H}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh}}\mathbf{w}_{\mathrm{P},k}^{\mathrm{h}}\mathbf{w}_{\mathrm{P},k}^{\mathrm{h},H}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh},H}\mathbf{v}_k^{\mathrm{h}},\\ b_{\mathrm{C},k}&=\mathbf{v}_k^{\mathrm{v},H}\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{vv}}\widetilde{\mathbf{d}}_{\mathrm{J},k}\widetilde{\mathbf{d}}_{\mathrm{J},k}^{H}\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{vv},H}\mathbf{v}_k^{\mathrm{v}},\\ b_{\mathrm{P},k}&=\mathbf{v}_k^{\mathrm{h},H}\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{hh}}\widetilde{\mathbf{d}}_{\mathrm{J},k}\widetilde{\mathbf{d}}_{\mathrm{J},k}^{H}\mathbf{\Xi}_{\mathrm{U},k}^{\mathrm{hh},H}\mathbf{v}_k^{\mathrm{h}},\\ z_{\mathrm{C},k}&=\mathbf{v}_k^{\mathrm{v},H}\mathbf{C}_{\mathrm{U}1,k}^{\mathrm{v}}\mathbf{v},z_{\mathrm{P},k}=\mathbf{v}_k^{\mathrm{h},H}\mathbf{C}_{\mathrm{U}1,k}^{\mathrm{h}}\mathbf{v}_k^{\mathrm{h}},\\ \overline{z}_{\mathrm{C},k}&=(z_{\mathrm{C},k}+a_{\mathrm{C},k})/b_{\mathrm{C},k}+p_{\mathrm{J}}^{\mathrm{v}},\widetilde{z}_{\mathrm{C},k}=z_{\mathrm{C},k}/b_{\mathrm{C},k}+p_{\mathrm{J}}^{\mathrm{v}},\\ \overline{z}_{\mathrm{P},k}&=(z_{\mathrm{P},k}+a_{\mathrm{P},k})/b_{\mathrm{P},k}+p_{\mathrm{J}}^{\mathrm{h}},\widetilde{z}_{\mathrm{P},k}=z_{\mathrm{P},k}/b_{\mathrm{P},k}+p_{\mathrm{J}}^{\mathrm{h}},\\ \overline{p}&=p_{\mathrm{J}}^{\mathrm{h}}-p_{\mathrm{J}}^{\mathrm{v}},\overline{z}_k=(\overline{z}_{\mathrm{P},k}-\overline{z}_{\mathrm{C},k})\,\overline{p},\widetilde{z}_k=(\widetilde{z}_{\mathrm{P},k}-\widetilde{z}_{\mathrm{C},k})\,\overline{p}.\end{aligned}$$

Proof: By utilizing dual-polarized channel model in (2), we can obtain the value of received jamming power at the different polarizations. Then, after analyzing the abovementioned values, we can easily obtain (11), which is omitted here for brevity. ■

B. Reduced Minorization-Maximization for $\mathbf{c}$, $\mathbf{w}_{\mathrm{C}}^{\mathrm{v}}$, and $\mathbf{w}_{\mathrm{P},k}^{\mathrm{h}}$

In this subsection, the optimization of the common rate portion vector $\mathbf{c}$ and digital precoder $\mathbf{w}_{\mathrm{C}}^{\mathrm{v}}$, $\mathbf{w}_{\mathrm{P},k}^{\mathrm{h}}$ are investigated, whose corresponding subproblem can be formulated as

$$\max_{\mathbf{V}_k,\mathbf{w}_{\mathrm{C}}^{\mathrm{v}},\mathbf{w}_{\mathrm{P},k}^{\mathrm{h}}}\ \min_{\bigcirc,\Delta}\sum\nolimits_{k=1}^{K}C_k+R_{\mathrm{P},k}\quad s.t.\ \mathrm{C1},\mathrm{C2},\mathrm{C3}.\tag{12}$$

However, as discussed before, problem (12) is challenging to solve due to the term $\min_{\bigcirc,\Delta}$, the non-convex expression of $R_{\mathrm{C},k}$ and $R_{\mathrm{P},k}$, the high-dimensional variables, and the multiple

constraints. Thus, in the following, we tackle the foregoing four challenges step by step. As for the term $\min_{\bigcirc,\Delta}$ inside (12), it can be removed by substituting ***Proposition 1*** and ***2*** into problem (12), resulting in the worst-case optimization subproblem (12).

Then, we propose the following detailed propositions for handling the high-dimensional variables.

Proposition 3 (Low-Dimensional Subspace Property for Dual-Polarized Networks): For any nontrivial stationary points $\mathbf{w}_{\mathrm{C}}^{\mathrm{v},\star}$ and $\mathbf{w}_{\mathrm{P},k}^{\mathrm{h},\star}$ in problem (12), they must lie in the range space spanned by their respective polarized channel's base $\mathbf{M}^{\mathrm{h},H}$ and $\mathbf{M}^{\mathrm{v},H}$, i.e., $\mathbf{w}_{\mathrm{C}}^{\mathrm{v},\star}=\mathbf{M}^{\mathrm{v},H}\mathbf{x}_{\mathrm{C}}^{\mathrm{v}}$ and $\mathbf{w}_{\mathrm{P},k}^{\mathrm{h},\star}=\mathbf{M}^{\mathrm{h},H}\mathbf{x}_{\mathrm{P},k}^{\mathrm{h}}$ with the unique $\mathbf{x}_{\mathrm{C}}^{\mathrm{v}}\in\mathbb{C}^{N_{\mathrm{U}}\times 1}$ and $\mathbf{x}_{\mathrm{P},k}^{\mathrm{h}}\in\mathbb{C}^{N_{\mathrm{U}}\times 1}$, where the full row-rank channel subspaces are given by

$$\mathbf{M}^{\mathrm{v}}=\left[\overline{\mathbf{\Omega}}_{\mathrm{U},1}^{\mathrm{vv},T},\cdots,\overline{\mathbf{\Omega}}_{\mathrm{U},K}^{\mathrm{vv},T}\right]^{T}\in\mathbb{C}^{N_{\mathrm{U}}\times N_{\mathrm{L}}},\tag{13}$$

$$\mathbf{M}^{\mathrm{h}}=\left[\overline{\mathbf{\Omega}}_{\mathrm{U},1}^{\mathrm{hh},T},\cdots,\overline{\mathbf{\Omega}}_{\mathrm{U},K}^{\mathrm{hh},T}\right]^{T}\in\mathbb{C}^{N_{\mathrm{U}}\times N_{\mathrm{L}}}.\tag{14}$$

Here, $N_{\mathrm{U}}=\sum_{k=1}^{K}N_{\mathrm{U},k}$.

Proof: The detailed proof is similar to those of [17] and thus is omitted in this paper for brevity. ■

By using ***Proposition 3***, the dimension of $\mathbf{w}_{\mathrm{C}}^{\mathrm{v}}$ and $\mathbf{w}_{\mathrm{P},k}^{\mathrm{h}}$ ($(K+1)N_{\mathrm{L}}$) can be greatly reduced to that of $\mathbf{x}_{\mathrm{C}}^{\mathrm{v}}$ and $\mathbf{x}_{\mathrm{P},k}^{\mathrm{h}}$ ($(K+1)N_{\mathrm{U}}$). Hence, (12) can be equivalently converted into

$$\max_{\mathbf{c},\mathbf{x}_{\mathrm{C}}^{\mathrm{v}},\mathbf{x}_{\mathrm{P},k}^{\mathrm{h}}}\sum\nolimits_{k=1}^{K}C_k+\widetilde{R}_{\mathrm{P},k}\quad s.t.\ \widetilde{\mathrm{C}}1,\widetilde{\mathrm{C}}2,\tag{15}$$

$$\widetilde{\mathrm{C}}3:\left\|\mathbf{M}^{\mathrm{v},H}\mathbf{x}_{\mathrm{C}}^{\mathrm{v}}\right\|^2+\sum\nolimits_{k=1}^{K}\left\|\mathbf{M}^{\mathrm{h},H}\mathbf{x}_{\mathrm{P},k}^{\mathrm{h}}\right\|^2\le P_{\max},$$

where $\widetilde{\mathrm{C}}$ denotes the modified constraint C with $\mathbf{x}_{\mathrm{C}}^{\mathrm{v}}$ and $\mathbf{x}_{\mathrm{P},k}^{\mathrm{h}}$. Although ***Proposition 3*** significantly reduces the dimensions of the optimization variables, problem (15) is challenging to solve due to the multiple constraints. Fortunately, we can find another important property to eliminate the power constraint $\widetilde{\mathrm{C}}3$.

Proposition 4 (Full Power Property): For the optimal point of (12), the total power constraint $\widetilde{\mathrm{C}}3$ must hold with equality.

Proof: The proposition can be proven immediately by using the power slackness condition of problem (12). ■

Building upon ***Proposition 4***, we can recast problem (15) as the following problem with reduced constraints:

$$\max_{\mathbf{c},\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v}},\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h}}}\sum\nolimits_{k=1}^{K}C_k+\overline{R}_{\mathrm{P},k}\quad s.t.\ \overline{\mathrm{C}}1,\overline{\mathrm{C}}2,\tag{16}$$

where $\overline{\mathrm{C}}$ is the modified version of $\widetilde{\mathrm{C}}$ by replacing $\mathbf{I}_{N_{\mathrm{U},k}}$ by $\left(\overline{P}_{\max}/P_{\max}\right)\mathbf{I}_{N_{\mathrm{U},k}}$. Here, $\overline{P}_{\max}$ is given by

$$\overline{P}_{\max}=\mathrm{Tr}\left(\overline{\mathbf{M}}^{\mathrm{v}}\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v}}\left(\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v}}\right)^{H}\right)+\sum\nolimits_{k=1}^{K}\mathrm{Tr}\left(\overline{\mathbf{M}}^{\mathrm{h}}\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h}}\left(\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h}}\right)^{H}\right),\tag{17}$$

where $\overline{\mathbf{M}}^{\mathrm{v}}=\mathbf{M}^{\mathrm{v}}\mathbf{M}^{\mathrm{v},H}$ and $\overline{\mathbf{M}}^{\mathrm{h}}=\mathbf{M}^{\mathrm{h}}\mathbf{M}^{\mathrm{h},H}$. As such, according to [17], after solving optimal $\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v},\star}$, $\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h},\star}$ in problem (16), we can obtain $\mathbf{w}_{\mathrm{C}}^{\mathrm{v},\star}$ and $\mathbf{w}_{\mathrm{P},k}^{\mathrm{h},\star}$ in (14) by calculating $\mathbf{w}_{\mathrm{C}}^{\mathrm{v},\star}=\sqrt{\kappa}\mathbf{M}^{\mathrm{v},H}\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v},\star}$ and $\mathbf{w}_{\mathrm{P},k}^{\mathrm{h},\star}=\sqrt{\kappa}\mathbf{M}^{\mathrm{h},H}\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h},\star}$, where $\kappa=P_{\max}/\overline{P}_{\max}^{\star}$ is a scaling factor.

Proposition 3 and ***Proposition 4*** suggest that problem (12) can be solved by handling problem (16) with reduced variables and constraints. However, the expressions for $R_{\mathrm{C},k}$ and $R_{\mathrm{P},k}$ are non-convex. Thus, the MM technique is adopted to convert $R_{\mathrm{C},k}$ and $R_{\mathrm{P},k}$ into solvable forms. Firstly, their lower bound are [11]

$$\begin{aligned}\overline{R}_{\mathrm{C},k}&\ge\log\det\left(\mathbf{S}^{H}\mathbf{O}_{\mathrm{C},k}^{(i_{\mathrm{d}}),-1}\mathbf{S}\right)-\mathrm{Tr}\left(\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}})}\left(\mathbf{O}_{\mathrm{C},k}-\mathbf{O}_{\mathrm{C},k}^{(i_{\mathrm{d}})}\right)\right),\\ \overline{R}_{\mathrm{P},k}&\ge\log\det\left(\mathbf{S}^{H}\mathbf{O}_{\mathrm{P},k}^{(i_{\mathrm{d}}),-1}\mathbf{S}\right)-\mathrm{Tr}\left(\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}})}\left(\mathbf{O}_{\mathrm{P},k}-\mathbf{O}_{\mathrm{P},k}^{(i_{\mathrm{d}})}\right)\right),\end{aligned}\tag{18}$$

where

$$\mathbf{O}_{\mathrm{C},k}=\begin{bmatrix}1 & \overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v},H}\mathbf{M}^{\mathrm{v}}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv},H}\\ \overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv}}\mathbf{M}^{\mathrm{v},H}\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v}} & \overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv}}\mathbf{M}^{\mathrm{v},H}\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v}}\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v},H}\mathbf{M}^{\mathrm{v}}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv},H}+\widetilde{\mathbf{C}}_{\mathrm{U},k}^{\mathrm{v}}\end{bmatrix},$$

$$\mathbf{O}_{\mathrm{P},k}=\begin{bmatrix}1 & \overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h},H}\mathbf{M}^{\mathrm{h}}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh},H}\\ \overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh}}\mathbf{M}^{\mathrm{h},H}\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h}} & \overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh}}\mathbf{M}^{\mathrm{h},H}\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h}}\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h},H}\mathbf{M}^{\mathrm{h}}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh},H}+\widetilde{\mathbf{C}}_{\mathrm{U},k}^{\mathrm{h}}\end{bmatrix},$$

$$\mathbf{R}_{\mathrm{C},k}=\mathbf{O}_{\mathrm{C},k}^{-1}\mathbf{S}\left(\mathbf{S}^{H}\mathbf{O}_{\mathrm{C},k}^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^{H}\mathbf{O}_{\mathrm{C},k}^{-1},\mathbf{S}=\left[\mathbf{I}_{N_{\mathrm{U},k}}\mathbf{0}_{N_{\mathrm{U},k}\times 1}\right]^{T},$$

$$\mathbf{R}_{\mathrm{P},k}=\mathbf{O}_{\mathrm{P},k}^{-1}\mathbf{S}\left(\mathbf{S}^{H}\mathbf{O}_{\mathrm{P},k}^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^{H}\mathbf{O}_{\mathrm{P},k}^{-1}$$

and $s^{(i_{\mathrm{d}})}$ is the solution obtained in the iteration i_{d}. Next, denoting $e_{\mathrm{C},k}=\log\det\left(\mathbf{S}^{H}\mathbf{O}_{\mathrm{C},k}^{(i_{\mathrm{d}}),-1}\mathbf{S}\right)+\mathrm{Tr}\left(\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}})}\mathbf{O}_{\mathrm{C},k}^{(i_{\mathrm{d}})}\right)$ and $e_{\mathrm{P},k}=\log\det\left(\mathbf{S}^{H}\mathbf{O}_{\mathrm{P},k}^{(i_{\mathrm{d}}),-1}\mathbf{S}\right)+\mathrm{Tr}\left(\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}})}\mathbf{O}_{\mathrm{P},k}^{(i_{\mathrm{d}})}\right)$, we obtain $\overline{R}_{\mathrm{C},k}\ge e_{\mathrm{C},k}-\mathrm{Tr}\left(\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}})}\mathbf{O}_{\mathrm{C},k}\right)$ and $\overline{R}_{\mathrm{P},k}\ge e_{\mathrm{P},k}-\mathrm{Tr}\left(\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}})}\mathbf{O}_{\mathrm{P},k}\right)$. Furthermore, $\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}})}$ and $\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}})}$ can be recast as

$$\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}})}=\begin{bmatrix}R_{\mathrm{C},k}^{(i_{\mathrm{d}}),11} & \mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),12}\\ \mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),21} & \mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),22}\end{bmatrix},\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}})}=\begin{bmatrix}R_{\mathrm{P},k}^{(i_{\mathrm{d}}),11} & \mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),12}\\ \mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),21} & \mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),22}\end{bmatrix},\tag{19}$$

where

$$\begin{aligned}R_{\mathrm{C},k}^{(i_{\mathrm{d}}),11}&=1+\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v},(i_{\mathrm{d}}),H}\mathbf{M}^{\mathrm{v}}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv},(i_{\mathrm{d}}),H}\widetilde{\mathbf{C}}_{\mathrm{U},k}^{\mathrm{v},(i_{\mathrm{d}}),-1}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv},(i_{\mathrm{d}})}\mathbf{M}^{\mathrm{v},H}\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v},(i_{\mathrm{d}})},\\ R_{\mathrm{P},k}^{(i_{\mathrm{d}}),11}&=1+\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h},(i_{\mathrm{d}}),H}\mathbf{M}^{\mathrm{h}}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh},(i_{\mathrm{d}}),H}\widetilde{\mathbf{C}}_{\mathrm{U},k}^{\mathrm{h},(i_{\mathrm{d}}),-1}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh},(i_{\mathrm{d}})}\mathbf{M}^{\mathrm{h},H}\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h},(i_{\mathrm{d}})},\\ \mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),12}&=-\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v},(i_{\mathrm{d}}),H}\mathbf{M}^{\mathrm{v}}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv},(i_{\mathrm{d}}),H}\widetilde{\mathbf{C}}_{\mathrm{U},k}^{\mathrm{v},(i_{\mathrm{d}}),-1},\\ \mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),12}&=-\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h},(i_{\mathrm{d}}),H}\mathbf{M}^{\mathrm{h}}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh},(i_{\mathrm{d}}),H}\widetilde{\mathbf{C}}_{\mathrm{U},k}^{\mathrm{h},(i_{\mathrm{d}}),-1},\\ \mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),21}&=\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),12,H},\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),22}=R_{\mathrm{C},k}^{(i_{\mathrm{d}}),11,-1}\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),21}\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),12},\\ \mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),21}&=\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),12,H},\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),22}=R_{\mathrm{P},k}^{(i_{\mathrm{d}}),11,-1}\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),21}\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),12}.\end{aligned}$$

Hence, $\mathrm{Tr}\left(\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}})}\mathbf{O}_{\mathrm{C},k}\right)$ and $\mathrm{Tr}\left(\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}})}\mathbf{O}_{\mathrm{P},k}\right)$ in $\overline{R}_{\mathrm{C},k}$ and $\overline{R}_{\mathrm{P},k}$ are transformed into

$$\begin{aligned}&\mathrm{Tr}\left(\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}})}\mathbf{O}_{\mathrm{C},k}\right)=R_{\mathrm{C},k}^{(i_{\mathrm{d}}),11}+2\Re\left(\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),12}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv}}\mathbf{M}^{\mathrm{v},H}\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v}}\right)\\ &+\mathrm{Tr}\left(\mathbf{R}_{\mathrm{C},k}^{(i_{\mathrm{d}}),22}\left(\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv}}\mathbf{M}^{\mathrm{v},H}\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v}}\overline{\mathbf{x}}_{\mathrm{C}}^{\mathrm{v},H}\mathbf{M}^{\mathrm{v}}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{vv},H}+\widetilde{\mathbf{C}}_{\mathrm{U},k}^{\mathrm{v}}\right)\right),\end{aligned}\tag{20}$$

$$\begin{aligned}&\mathrm{Tr}\left(\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}})}\mathbf{O}_{\mathrm{P},k}\right)=R_{\mathrm{P},k}^{(i_{\mathrm{d}}),11}+2\Re\left(\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),12}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh}}\mathbf{M}^{\mathrm{h},H}\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h}}\right)\\ &+\mathrm{Tr}\left(\mathbf{R}_{\mathrm{P},k}^{(i_{\mathrm{d}}),22}\left(\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh}}\mathbf{M}^{\mathrm{h},H}\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h}}\overline{\mathbf{x}}_{\mathrm{P},k}^{\mathrm{h},H}\mathbf{M}^{\mathrm{h}}\overline{\mathbf{\Omega}}_{\mathrm{U},k}^{\mathrm{hh},H}+\widetilde{\mathbf{C}}_{\mathrm{U},k}^{\mathrm{h}}\right)\right).\end{aligned}\tag{21}$$

By substituting (18)-(21) into problem (16) and dropping the constant terms, we can rewrite problem (16) as

$$\begin{aligned}\min_{\mathbf{c},\overline{\mathbf{x}}^{\rm v}_{\rm C},\overline{\mathbf{x}}^{\rm h}_{{\rm P},k}} \; & -C+\sum\nolimits_{k=1}^{K}\Big(2\Re\left(\mathbf{R}^{(i_{\rm d}),12}_{{\rm P},k}\overline{\mathbf{\Omega}}^{\rm hh}_{{\rm U},k}\mathbf{M}^{{\rm h},H}\overline{\mathbf{x}}^{\rm h}_{{\rm P},k}\right)\\ &+\sum\nolimits_{m=1}^{M}\overline{\mathbf{x}}^{{\rm h},H}_{{\rm P},m}\mathbf{M}^{\rm h}\overline{\mathbf{\Omega}}^{{\rm hh},H}_{{\rm U},k}\mathbf{R}^{(i_{\rm d}),22}_{{\rm P},k}\overline{\mathbf{\Omega}}^{\rm hh}_{{\rm U},k}\mathbf{M}^{{\rm h},H}\overline{\mathbf{x}}^{\rm h}_{{\rm P},m}\\ &+\overline{\mathbf{x}}^{{\rm v},H}_{\rm C}\mathbf{M}^{\rm v}\overline{\mathbf{\Omega}}^{{\rm hv},H}_{{\rm U},k}\mathbf{R}^{(i_{\rm d}),22}_{{\rm P},k}\overline{\mathbf{\Omega}}^{\rm hv}_{{\rm U},k}\mathbf{M}^{{\rm v},H}\overline{\mathbf{x}}^{\rm v}_{\rm C}\Big) \qquad (22)\\ s.t.\ \overline{\underline{\rm C}}1: & \sum\nolimits_{m=1}^{M}\overline{\mathbf{x}}^{{\rm h},H}_{{\rm P},m}\mathbf{M}^{\rm h}\overline{\mathbf{\Omega}}^{{\rm hh},H}_{{\rm U},k}\mathbf{R}^{(i_{\rm d}),22}_{{\rm P},k}\overline{\mathbf{\Omega}}^{\rm hh}_{{\rm U},k}\mathbf{M}^{{\rm h},H}\overline{\mathbf{x}}^{\rm h}_{{\rm P},m}\\ &+\overline{\mathbf{x}}^{{\rm v},H}_{\rm C}\mathbf{M}^{\rm v}\overline{\mathbf{\Omega}}^{{\rm hv},H}_{{\rm U},k}\mathbf{R}^{(i_{\rm d}),22}_{{\rm P},k}\overline{\mathbf{\Omega}}^{\rm hv}_{{\rm U},k}\mathbf{M}^{{\rm v},H}\overline{\mathbf{x}}^{\rm v}_{\rm C}\\ &+2\Re\left(\mathbf{R}^{(i_{\rm d}),12}_{{\rm P},k}\overline{\mathbf{\Omega}}^{\rm hh}_{{\rm U},k}\mathbf{M}^{{\rm h},H}\overline{\mathbf{x}}^{\rm h}_{{\rm P},k}\right)\le\widetilde{R}_k,\forall k,\\ \overline{\underline{\rm C}}2: & \overline{\mathbf{x}}^{{\rm v},H}_{\rm C}\mathbf{M}^{\rm v}\overline{\mathbf{\Omega}}^{{\rm vv},H}_{{\rm U},k}\mathbf{R}^{(i_{\rm d}),22}_{{\rm C},k}\overline{\mathbf{\Omega}}^{\rm vv}_{{\rm U},k}\mathbf{M}^{{\rm v},H}\overline{\mathbf{x}}^{\rm v}_{\rm C}\\ &+\sum\nolimits_{m=1}^{M}\overline{\mathbf{x}}^{{\rm h},H}_{{\rm P},m}\mathbf{M}^{\rm h}\overline{\mathbf{\Omega}}^{{\rm vh},H}_{{\rm U},k}\mathbf{R}^{(i_{\rm d}),22}_{{\rm C},k}\overline{\mathbf{\Omega}}^{\rm vh}_{{\rm U},k}\mathbf{M}^{{\rm h},H}\overline{\mathbf{x}}^{\rm h}_{{\rm P},m}\\ &+2\Re\left(\mathbf{R}^{(i_{\rm d}),12}_{{\rm C},k}\overline{\mathbf{\Omega}}^{\rm vv}_{{\rm U},k}\mathbf{M}^{{\rm v},H}\overline{\mathbf{x}}^{\rm v}_{\rm C}\right)\le\widetilde{C}_k,\forall k,\end{aligned}$$

where $C=\sum_{k=1}^{K}C_k$,

$$\widetilde{C}_k=-{\rm Tr}\left(\mathbf{R}^{(i_{\rm d}),22}_{{\rm C},k}\left(\sigma^{{\rm v},2}_k\mathbf{I}_{N_{{\rm U},k}}+\widetilde{\mathbf{C}}^{\rm v}_{{\rm J},k}\right)\right)-C+e_{{\rm C},k}-R^{(i_{\rm d}),11}_{{\rm C},k},$$

$$\widetilde{R}_k=-{\rm Tr}\left(\mathbf{R}^{(i_{\rm d}),22}_{{\rm P},k}\left(\sigma^{{\rm h},2}_k\mathbf{I}_{N_{{\rm U},k}}+\widetilde{\mathbf{C}}^{\rm h}_{{\rm J},k}\right)\right)-\overline{R}_{k,\min}+e_{{\rm P},k}-R^{(i_{\rm d}),11}_{{\rm P},k}.$$

Note that (22) is a QCQP with low-dimensional variables and reduced constraints, which can be solved using the CVX optimization toolbox efficiently. The R-MM algorithm can converge to stationary points, where the final solutions are calculated as $\mathbf{w}^{\rm v}_{\rm C}=\sqrt{\kappa}\mathbf{M}^{{\rm v},H}\overline{\mathbf{x}}^{\rm v}_{\rm C}$ and $\mathbf{w}^{\rm h}_{{\rm P},k}=\sqrt{\kappa}\mathbf{M}^{{\rm h},H}\overline{\mathbf{x}}^{\rm h}_{{\rm P},k}$.

C. CCD-Based K-Bisection BSUM for $\mathbf{P}_{{\rm L},a}$ and $\mathbf{P}_{{\rm U},kr}$

This subsection first describes the design of the T-RIS's phase-shift matrix $\mathbf{P}_{{\rm L},a}$. Denoting $\mathbf{W}^{\rm v}_{\rm C}=\mathbf{w}^{\rm v}_{\rm C}\mathbf{w}^{{\rm v},H}_{\rm C}$ and $\mathbf{W}^{\rm h}_{\rm P}=\sum_{m=1}^{M}\mathbf{w}^{\rm h}_{{\rm P},m}\mathbf{w}^{{\rm h},H}_{{\rm P},m}$, the subproblem for $\mathbf{P}_{{\rm L},a}$ is given by

$$\begin{aligned}\min_{\mathbf{P}_{{\rm L},a}} & \sum\nolimits_{k=1}^{K}\Big(2\Re\left(\mathbf{R}^{(i_{\rm d}),12}_{{\rm P},k}\overline{\mathbf{\Omega}}^{\rm hh}_{{\rm U},k}\mathbf{w}^{\rm h}_{{\rm P},k}\right) \qquad (23)\\ &+{\rm Tr}\left(\mathbf{R}^{(i_{\rm d}),22}_{{\rm P},k}\left(\overline{\mathbf{\Omega}}^{\rm hh}_{{\rm U},k}\mathbf{W}^{\rm h}_{\rm P}\overline{\mathbf{\Omega}}^{{\rm hh},H}_{{\rm U},k}+\overline{\mathbf{\Omega}}^{\rm hv}_{{\rm U},k}\mathbf{W}^{\rm v}_{\rm C}\overline{\mathbf{\Omega}}^{{\rm hv},H}_{{\rm U},k}\right)\right)\Big)\\ s.t.\ \underline{\rm C}1: & {\rm Tr}\left(\mathbf{R}^{(i_{\rm d}),22}_{{\rm P},k}\left(\overline{\mathbf{\Omega}}^{\rm hh}_{{\rm U},k}\mathbf{W}^{\rm h}_{\rm P}\overline{\mathbf{\Omega}}^{{\rm hh},H}_{{\rm U},k}+\overline{\mathbf{\Omega}}^{\rm hv}_{{\rm U},k}\mathbf{W}^{\rm v}_{\rm C}\overline{\mathbf{\Omega}}^{{\rm hv},H}_{{\rm U},k}\right)\right)\\ &+2\Re\left(\mathbf{R}^{(i_{\rm d}),12}_{{\rm P},k}\overline{\mathbf{\Omega}}^{\rm hh}_{{\rm U},k}\mathbf{w}^{\rm h}_{{\rm P},k}\right)+\le\widetilde{R}_k,\forall k,\\ \underline{\rm C}2: & {\rm Tr}\left(\mathbf{R}^{(i_{\rm d}),22}_{{\rm C},k}\left(\overline{\mathbf{\Omega}}^{\rm vv}_{{\rm U},k}\mathbf{W}^{\rm v}_{\rm C}\overline{\mathbf{\Omega}}^{{\rm vv},H}_{{\rm U},k}+\overline{\mathbf{\Omega}}^{\rm vh}_{{\rm U},k}\mathbf{W}^{\rm h}_{\rm P}\overline{\mathbf{\Omega}}^{{\rm vh},H}_{{\rm U},k}\right)\right)\\ &+2\Re\left(\mathbf{R}^{(i_{\rm d}),12}_{{\rm C},k}\overline{\mathbf{\Omega}}^{\rm vv}_{{\rm U},k}\mathbf{w}^{\rm v}_{\rm C}\right)+\le\widetilde{C}_k,\forall k,{\rm C}4.\end{aligned}$$

By using ***Lemma*** in [11] and some matrix transformations, the subproblem to $\mathbf{p}_{{\rm L},a}$ can be recast as

$$\begin{aligned}\min_{\mathbf{p}_{{\rm L},a}} & \sum\nolimits_{k=1}^{K}\left(\mathbf{p}^H_{{\rm L},a}\overline{\mathbf{U}}^{\rm h}_{{\rm L},ka}\mathbf{p}_{{\rm L},a}+2\Re\left(\mathbf{p}^H_{{\rm L},a}\mathbf{e}^{\rm h}_{{\rm L},ka}\right)\right) \qquad (24)\\ s.t.\ \underline{\underline{\rm C}}1: & g_{1,k}\left(\mathbf{p}_{{\rm L},a}\right)=\mathbf{p}^H_{{\rm L},a}\overline{\mathbf{U}}^{\rm h}_{{\rm L},ka}\mathbf{p}_{{\rm L},a}\\ &\qquad+2\Re\left(\mathbf{p}^H_{{\rm L},a}\mathbf{e}^{\rm h}_{{\rm L},ka}\right)-\widetilde{R}_k\le0,\forall k,\\ \underline{\underline{\rm C}}2: & g_{2,k}\left(\mathbf{p}_{{\rm L},a}\right)=\mathbf{p}^H_{{\rm L},a}\overline{\mathbf{U}}^{\rm v}_{{\rm L},ka}\mathbf{p}_{{\rm L},a}\\ &\qquad+2\Re\left(\mathbf{p}^H_{{\rm L},a}\mathbf{e}^{\rm v}_{{\rm L},ka}\right)-\widetilde{C}_k\le0,\forall k,\\ \underline{\underline{\rm C}}4: & \left|\left[\mathbf{p}_{{\rm L},a}\right]_n\right|=1,\forall n,a,\end{aligned}$$

where

$$\mathbf{p}_{{\rm L},a}=\begin{bmatrix}\mathbf{p}^{\rm vv}_{{\rm L},a}\\ \mathbf{p}^{\rm hh}_{{\rm L},a}\end{bmatrix},\mathbf{e}^{\rm v}_{{\rm L},ka}=\begin{bmatrix}\mathbf{e}^{{\rm vv},*}_{{\rm L},ka}\\ \mathbf{0}_{M_{{\rm L},a}\times1}\end{bmatrix},$$

$$\overline{\mathbf{U}}^{\rm v}_{{\rm L},ka}=\begin{bmatrix}\mathbf{U}^{\rm vv}_{{\rm L},ka} & \\ & \mathbf{U}^{\rm vh}_{{\rm L},ka}\end{bmatrix},$$

$$\overline{\mathbf{U}}^{\rm h}_{{\rm L},ka}=\begin{bmatrix}\mathbf{U}^{\rm hv}_{{\rm L},ka} & \\ & \mathbf{U}^{\rm hh}_{{\rm L},ka}\end{bmatrix},\mathbf{e}^{\rm h}_{{\rm L},ka}=\begin{bmatrix}\mathbf{0}_{M_{{\rm L},a}\times1}\\ \mathbf{e}^{{\rm hh},*}_{{\rm L},ka}\end{bmatrix},$$

$$\mathbf{\Omega}^{ii}_{{\rm L},(u,v)}=\begin{cases}\prod_{a=1}^{A}\mathbf{P}^{ii}_{{\rm L},a}\mathbf{B}^{ii}_{{\rm L},a},u,v\in[A]\\ \mathbf{I}_{N_{\rm L}},u=1,v=0,\\ \mathbf{I}_{M_{{\rm L},A}},u=A+1,v=A,\end{cases},ii=\{{\rm hh},{\rm vv}\},$$

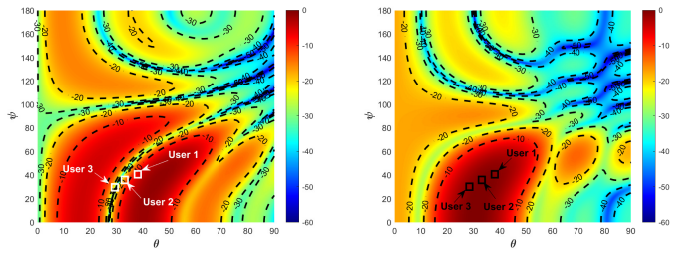
$$\begin{aligned}&\mathbf{e}^{\rm hh}_{{\rm L},ka}\\ &={\rm diag}\Big\{\mathbf{B}^{\rm hh}_{{\rm L},a}\mathbf{\Omega}^{\rm hh}_{{\rm L},(1,a-1)}\mathbf{w}^{\rm h}_{{\rm P},k}\mathbf{R}^{(i_{\rm d}),12}_{{\rm P},k}\mathbf{\Xi}^{\rm hh}_{{\rm U},k}\mathbf{D}^{\rm hh}_{{\rm U},k}\mathbf{\Omega}^{\rm hh}_{{\rm L},(a+1,A)}\Big\},\\ &\mathbf{e}^{\rm vv}_{{\rm L},ka}\\ &={\rm diag}\left\{\mathbf{B}^{\rm vv}_{{\rm L},a}\mathbf{\Omega}^{\rm vv}_{{\rm L},(1,a-1)}\mathbf{w}^{\rm v}_{\rm C}\mathbf{R}^{(i_{\rm d}),12}_{{\rm C},k}\mathbf{\Xi}^{\rm vv}_{{\rm U},k}\mathbf{D}^{\rm vv}_{{\rm U},k}\mathbf{\Omega}^{\rm vv}_{{\rm L},(a+1,A)}\right\},\\ &\mathbf{U}^{\rm hh}_{{\rm L},ka}\\ &=\Big(\mathbf{\Omega}^{{\rm hh},H}_{{\rm L},(a+1,A)}\mathbf{D}^{{\rm hh},H}_{{\rm U},k}\mathbf{\Xi}^{{\rm hh},H}_{{\rm U},k}\mathbf{R}^{(i_{\rm d}),22}_{{\rm P},k}\mathbf{\Xi}^{\rm hh}_{{\rm U},k}\mathbf{D}^{\rm hh}_{{\rm U},k}\mathbf{\Omega}^{\rm hh}_{{\rm L},(a+1,A)}\Big)\\ &\qquad\odot\left(\mathbf{B}^{\rm hh}_{{\rm L},a}\mathbf{\Omega}^{\rm hh}_{{\rm L},(1,a-1)}\mathbf{W}^{\rm h}_{\rm P}\left(\mathbf{B}^{\rm hh}_{{\rm L},a}\mathbf{\Omega}^{\rm hh}_{{\rm L},(1,a-1)}\right)^H\right)^T,\\ &\mathbf{U}^{\rm hv}_{{\rm L},ka}\\ &=\Big(\mathbf{\Omega}^{{\rm vv},H}_{{\rm L},(a+1,A)}\mathbf{D}^{{\rm hv},H}_{{\rm U},k}\mathbf{\Xi}^{{\rm hh},H}_{{\rm U},k}\mathbf{R}^{(i_{\rm d}),22}_{{\rm P},k}\mathbf{\Xi}^{\rm hh}_{{\rm U},k}\mathbf{D}^{\rm hv}_{{\rm U},k}\mathbf{\Omega}^{\rm vv}_{{\rm L},(a+1,A)}\Big)\\ &\qquad\odot\left(\mathbf{B}^{\rm vv}_{{\rm L},a}\mathbf{\Omega}^{\rm vv}_{{\rm L},(1,a-1)}\mathbf{W}^{\rm v}_{\rm C}\left(\mathbf{B}^{\rm vv}_{{\rm L},a}\mathbf{\Omega}^{\rm vv}_{{\rm L},(1,a-1)}\right)^H\right)^T\\ &\mathbf{U}^{\rm vv}_{{\rm L},ka}\\ &=\Big(\mathbf{\Omega}^{{\rm vv},H}_{{\rm L},(a+1,A)}\mathbf{D}^{{\rm vv},H}_{{\rm U},k}\mathbf{\Xi}^{{\rm vv},H}_{{\rm U},k}\mathbf{R}^{(i_{\rm d}),22}_{{\rm C},k}\mathbf{\Xi}^{\rm vv}_{{\rm U},k}\mathbf{D}^{\rm vv}_{{\rm U},k}\mathbf{\Omega}^{\rm vv}_{{\rm L},(a+1,A)}\Big)\\ &\qquad\odot\left(\mathbf{B}^{\rm vv}_{{\rm L},a}\mathbf{\Omega}^{\rm vv}_{{\rm L},(1,a-1)}\mathbf{W}^{\rm v}_{\rm C}\left(\mathbf{B}^{\rm vv}_{{\rm L},a}\mathbf{\Omega}^{\rm vv}_{{\rm L},(1,a-1)}\right)^H\right)^T,\\ &\mathbf{U}^{\rm vh}_{{\rm L},ka}\\ &=\Big(\mathbf{\Omega}^{{\rm hh},H}_{{\rm L},(a+1,A)}\mathbf{D}^{{\rm vh},H}_{{\rm U},k}\mathbf{\Xi}^{{\rm vv},H}_{{\rm U},k}\mathbf{R}^{(i_{\rm d}),22}_{{\rm C},k}\mathbf{\Xi}^{\rm vv}_{{\rm U},k}\mathbf{D}^{\rm vh}_{{\rm U},k}\mathbf{\Omega}^{\rm hh}_{{\rm L},(a+1,A)}\Big)\\ &\qquad\odot\left(\mathbf{B}^{\rm hh}_{{\rm L},a}\mathbf{\Omega}^{\rm hh}_{{\rm L},(1,a-1)}\mathbf{W}^{\rm h}_{\rm P}\left(\mathbf{B}^{\rm hh}_{{\rm L},a}\mathbf{\Omega}^{\rm hh}_{{\rm L},(1,a-1)}\right)^H\right)^T.\end{aligned}$$

Nevertheless, problem (24) is also intractable due to the high-dimensional dual-polarized variables, and the multiple QoS constraints with UMC. Thus, the K-bisection BSUM is proposed for handling the above two challenges. Specifically, by adding $\underline{\underline{\rm C}}1$ and $\underline{\underline{\rm C}}2$ into the objective function in (24) with non-negative Lagrange dual variables, the BSUM in [5] is first adopted to solve $\mathbf{p}_{{\rm L},a}$, and then K-bisection is proposed to obtain the optimal Lagrange dual variables, as detailed next.

Firstly, by introducing ϖ_k and ς_k, the Lagrange dual problem of problem (24) can be expressed as

$$\begin{aligned}\min_{\mathbf{p}_{{\rm L},a}} & \, h\left(\mathbf{p}_{{\rm L},a}\right)=\mathbf{p}^H_{{\rm L},a}\widetilde{\mathbf{U}}_{{\rm L},a}\mathbf{p}_{{\rm L},a}+2\Re\left(\mathbf{p}^H_{{\rm L},a}\widetilde{\mathbf{e}}_{{\rm L},a}\right)\\ s.t. & \ \ \underline{\underline{\rm C}}4: \left|\left[\mathbf{p}_{{\rm L},a}\right]_n\right|=1,\forall n,a, \qquad (25)\end{aligned}$$

where $\widetilde{\mathbf{U}}_{{\rm L},a}=\sum_{k=1}^{K}\left(1+\varpi_k\right)\overline{\mathbf{U}}^{\rm h}_{{\rm L},ka}+\varsigma_k\overline{\mathbf{U}}^{\rm v}_{{\rm L},ka}$ and $\widetilde{\mathbf{e}}_{{\rm L},a}=\sum_{k=1}^{K}\left(1+\varpi_k\right)\mathbf{e}^{\rm h}_{{\rm L},ka}+\varsigma_k\mathbf{e}^{\rm v}_{{\rm L},ka}$. Then, BSUM using the cyclic coordinate descent (CCD) algorithm is proposed, where the optimization variables $\mathbf{p}_{{\rm L},a}$ are divided into $\overline{M}_{{\rm L},a}=2M_{{\rm L},a}$



(a) (b)

Fig. 2: (a) The private beampattern in polarization h; (b) The common beampattern in polarization v.

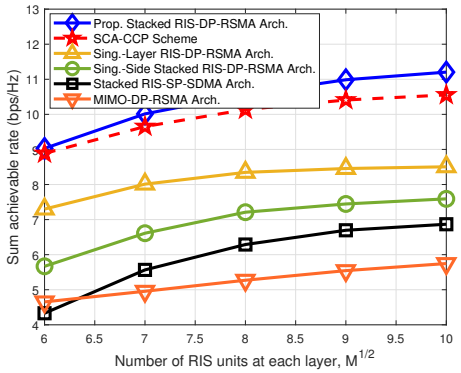


Fig. 3: Sum achievable rate versus number of RIS units $\sqrt{M}$.

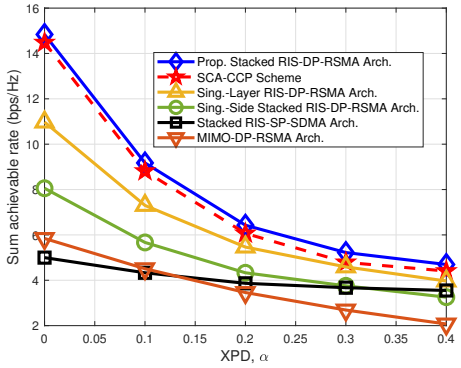


Fig. 4: Sum achievable rate versus XPD α.

subproblems, and then addressed in a block-by-block manner. To elaborate, we only solve the subproblem w.r.t. $p_{\mathrm{L},ai}$ in each block, while the the remaining variables are updated block-wise such that the closed-form solutions can be obtained, i.e., [5]

$$p_{\mathrm{L},ai} = e^{\left\{ j \arg\left(-\sum_{j=1}^{j<i} \widetilde{U}_{\mathrm{L},a}^{(i,j)} p_{\mathrm{L},aj}^{(i_\mathrm{d})} - \sum_{j>1}^{M_{\mathrm{L},a}} \widetilde{U}_{\mathrm{L},a}^{(i,j)} p_{\mathrm{L},aj}^{(i_\mathrm{d})} - 2\widetilde{e}_{\mathrm{L},ai} \right) \right\}}. \tag{26}$$

However, the proposed BSUM is executed with the fixed Lagrange dual variables ϖ_k and ς_k. Thus, the K-bisection method is adopted to search for the optimal ϖ_k and ς_k [5]. Under the iterative framework, $\mathbf{p}_{\mathrm{L},a}$, ϖ_k, and ς_k can be optimized until converging to stationary points.

Next, we turn to the optimization of $\mathbf{P}_{\mathrm{U},kr}$. Note that CCD-Based K-Bisection BSUM is also applicable to the design of $\mathbf{P}_{\mathrm{U},kr}$ by using the similar matrix transformations, and thus is omitted here due to the space limit.

IV. Simulation Results

This section provides numerical results to evaluate the validity and superiority of our proposed architecture and optimization framework. We consider a scenario where UAV equipped with $N_\mathrm{L} = 6 \times 6$ Tx antennas and A layers of T-RIS serves $K = 3$ users with $R = 2$ layers of R-RIS and $N_{\mathrm{U},k} = 2 \times 2$ Rx antennas, in the presence of one jammer, where the altitudes of UAV and jammer are set as 60 m and 30 m, respectively. The noise powers at the user's the vertical and horizontal polarization are set as $\sigma_k^{\mathrm{v},2} = \sigma_k^{\mathrm{h},2} = -60$ dBm, the carrier frequency is 5.8 GHz, and the maximum transmit power of UAV is $P_{\max} = 40$ dBm. In addition, the total jamming power is set as $P_{\mathrm{J},\max} = P_{\max} + 10\log(\mathrm{SJNR})$ [dBm] and the jammer divides the total power into two parts in vertical and horizontal polarization randomly, where SJNR denotes the signal-to-jamming plus noise-ratio (SJNR). The achievable rate target is set as $R_{k,\min} = 1$ bps/Hz, and the XPD is $\alpha_{\mathrm{U},k} = \alpha_{\mathrm{J},k} = 0.1$. Moreover, the CSI uncertainty is defined as $\Delta = \theta_U - \theta_L = \varphi_U - \varphi_L$.

Fig. 2 illustrates the normalized beampattern in different polarizations, where $\Delta = 4^\circ$ and $\mathrm{SJNR} = -30$ dB. We observe that the private beampatterns in polarization h can accurately align their mainlobes with the desired target (i.e., user 1), even with Δ. However, due to users' overlapping channel angular sectors, the nulls can only be generated at the region of undesired user having non-overlapping sectors (user 3), which results in severe multi-user interference for the remaining user (user 2). Fortunately, by utilization of dual-polarized RSMA, the common beampatterns in polarization v can align their mainlobes with all the users for transmitting the common message, leading to the effective suppression of the multi-user interference.

Fig. 3 illustrates the sum achievable rate versus the number of RIS units at each layer $\sqrt{M}$. Here, we set $M_{\mathrm{L},a} = M_{\mathrm{U},kr} = M$ and SJNR= -10 dB. It can be observed that for all these architectures and schemes, the sum achievable rate increases with M due to the increasing aperture gain of the transceiver. In addition, the proposed stacked RIS-DP-RSMA architecture achieves the highest achievable rate. The reason is the fact that the amplitudes of the desired and jamming signals impinging on the stacked RIS-aided transceiver can be enhanced constructively and weakened destructively in a larger range, respectively, than the benchmark architectures. Besides, by using the dual-polarized RSMA, not only the inherent cross-polarization interference between the common message and private one but also the inter-user interference can be mitigated in the power and polarization domains, resulting in the significant performance enhancement of our proposed architecture compared to the stacked RIS-SP-SDMA architecture.

Fig. 4 presents the impact of XPD on the sum achievable rate. Clearly, due to the increased cross-polarization interference, the sum achievable rate of all DP-RSMA architectures decreases sharply with the increasing α, especially when $0 \le \alpha \le 0.2$. More specifically, as $\alpha \to 0$, the user can recover common/private messages without interference of private/common messages, i.e., in ideal conditions, resulting in the performance enhancement. In addition, the SP-SDMA only experiences the fading effects of α on one specific polarization instead of the cross-polarization interference, and thus its sum achievable rate

is not sensitive to α, which leads to the higher sum achievable rate than that of MIMO-DP-RSMA when $\alpha \geq 0.2$.

V. Conclusions

This paper proposed a novel architecture of stacked RIS-aided dual-polarized UAV-RSMA networks, which can fully unleash the potentials of dual polarization, RSMA, and stacked RIS-transceiver in the power, space, and polarization domains to mitigate the inter-user interference and suppress the malicious jamming attacks. Besides, we formulated a general sum rate maximization problem with the jammer's imperfect CSI and unknown XPD. To handle the formulated intractable problem, a novel low-complexity optimization framework was developed by using quadratic property, R-MM algorithm, and CCD-based K-bisection BSUM, which provides semi-closed-form solutions. The simulation results confirmed the superiority of our proposed architecture and framework in comparison with the existing ones.

References

[1] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, 2016.

[2] A. Mukherjee, S. A. A. Fakoorian, and *et.al.*, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, 2014.

[3] B. Clerckx and *et.al.*, "A primer on rate-splitting multiple access: Tutorial, myths, and frequently asked questions," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 5, pp. 1265–1308, 2023.

[4] H. Wang, L. Zhang, T. Li, and J. Tugnait, "Spectrally efficient jamming mitigation based on code-controlled frequency hopping," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 728–732, 2011.

[5] Y. Sun and *et al.*, "Active-passive cascaded RIS-aided receiver design for jamming nulling and signal enhancing," *IEEE Trans. Wire. Commun., Early Access*, 2023.

[6] A. Mishra, Y. Mao, O. Dizdar, and B. Clerckx, "Rate-splitting multiple access for 6G–Part I: Principles, applications and future works," *IEEE Commun. Letters*, vol. 26, no. 10, pp. 2232–2236, 2022.

[7] Z. Lin and *et al.*, "Supporting iot with rate-splitting multiple access in satellite and aerial-integrated networks," *IEEE Int. Things J.*, vol. 8, no. 14, pp. 11 123–11 134, 2021.

[8] E. Piovano and *et al.*, "Optimal DoF region of the K-user MISO BC with partial CSIT," *IEEE Commun. Lett.*, vol. 21, no. 11, pp. 2368–2371, 2017.

[9] H. Li, Y. Mao, O. Dizdar, and B. Clerckx, "Rate-splitting multiple access for 6G–Part III: Interplay with reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, vol. 26, no. 10, pp. 2242–2246, 2022.

[10] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2020.

[11] Y. Sun and *et al.*, "RIS-assisted robust hybrid beamforming against simultaneous jamming and eavesdropping attacks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 11, pp. 9212–9231, 2022.

[12] S. Li and *et.al.*, "Reconfigurable intelligent surface assisted UAV communication: Joint trajectory design and passive beamforming," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 716–720, 2020.

[13] Y. Sun and *et.al.*, "Joint transmissive and reflective RIS-aided secure MIMO systems design under spatially-correlated angular uncertainty and coupled PSEs," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 3606–3621, 2023.

[14] Y. Sun and *et al.*, "Energy-efficient hybrid beamforming for multilayer RIS-assisted secure integrated terrestrial-aerial networks," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4189–4210, 2022.

[15] A. S. de Sena and *et.al.*, "Dual-polarized massive MIMO-RSMA networks: Tackling imperfect SIC," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 3194–3215, 2023.

[16] R. Ma, W. Yang, X. Guan, X. Lu, Y. Song, and D. Chen, "Covert mmwave communications with finite blocklength against spatially random wardens," *IEEE Int. Things J.*, vol. 11, no. 2, pp. 3402–3416, 2024.

[17] X. Zhao, S. Lu, Q. Shi, and Z.-Q. Luo, "Rethinking WMMSE: Can its complexity scale linearly with the number of BS antennas?" *IEEE Trans. Signal Process.*, vol. 71, pp. 433–446, 2023.